

Figure 1A

1	CTGCGACCGGTCTGCCACTTGCCAGGTGACCGCTGATGGGAAGACCAGCTGTGTGTGCAG	60
61	GGAAAGCGAGGTGGGGGATGGGCGTGCCCTGCTACGGACACCTGCTCCACGAGGTGCAGAA	120
121	GGCCACGCAGACAGGCCGGGTGTTCTGTCAGCTGAGGGTCGCCGTGGCCATGATGGACCA	180
1		M M D Q 4
181	GGGCTGCCGGGAAATCCTTACCACAGCGGGCCCTTTCACCGTGCTGGTGCCATCCGTCTC	240
5	G C R E I L T T A G P F T V L V P S V S	24
241	CTCCTTCTCCTCCAGGACCATGAATGCATCCCTTGCCCAGCAGCTCTGTAGACAGCACAT	300
25	S F S S R T M N A S L A Q Q L C R Q H I	44
301	CATCGCAGGGCAGCACATCCTGGAGGACACAAGGACCCAACAAACACGAAGGTGGTGGAC	360
45	I A G Q H I L E D T R T Q Q T R R W W T	64
361	GCTGGCCGGGCAGGAGATCACCGTCACCTTTAACCAATTTCACGAAATACTCCTACAAGTA	420
65	L A G Q E I T V T F N Q F T K Y S Y K Y	84
421	CAAAGACCAGCCCCAGCAGACGTTCAACATCTACAAGGCCAACACATAGCAGCTAATGG	480
85	K D Q P Q Q T F N I Y K A N N I A A N G	104
481	CGTCTTCCACGTGGTCACTGGCCTGCGGTGGCAGGCCCCCTCTGGGACCCCTGGGGATCC	540
105	V F H V V T G L R W Q A P S G T P G D P	124
541	CAAGAGAACTATCGGACAGATCCTCGCCTCTACCGAGGCCTTCAGCCGCTTTGAAACCAT	600
125	K R T I G Q I L A S T E A F S R F E T I	144
601	CCTGGAGAACTGTGGGCTGCCCTCCATCCTGGACGGACCTGGGCCCTTCACAGTCTTTGC	660
145	L E N C G L P S I L D G P G P F T V F A	164
661	CCCAAGCAATGAGGCTGTGGACAGCTTGCGTGACGGCCGCTGATCTACCTCTTCACAGC	720
165	P S N E A V D S L R D G R L I Y L F T A	184
721	GGGTCTCTCTAAACTGCAGGAGTTGGTGCGGTACCACATCTACAACCACGGCCAGCTGAC	780
185	G L S K L Q E L V R Y H I Y N H G Q L T	204
781	CGTTGAGAAGCTCATCTCAAGGGTCGGATCCTCACCATGGCGAACCAGGTCTGGCTGT	840
205	V E K L I S K G R I L T M A N Q V L A V	224

Figure 1B

841	GAACATTTCTGAGGAGGGGCGCATCCTGCTGGGACCCGAGGGGTCCCGCTGCAGAGGGT	900
225	N I S E E G R I L L G P E G V P L Q R V	244
901	AGACGTGATGGCCGCAATGGCGTGATCCACATGCTGGACGGCATCCTGCTGCCCCGAC	960
245	D V M A A N G V I H M L D G I L L P P T	264
961	CATCCTGCCCATCCTGCCAAGCACTGCAGCGAGGAGCAGCACAAGATTGTGGCGGGCTC	1020
265	I L P I L P K H C S E E Q H K I V A G S	284
1021	CTGTGTGGACTGCCAAGCCCTGAACACCAGCACGTGTCCCCCAACAGTGTGAAGCTGGA	1080
285	C V D C Q A L N T S T C P P N S V K L D	304
1081	CATCTTCCCAAGGAGTGTGTCTACATCCATGACCCAACGGGGCTCAATGTGCTAAAGAA	1140
305	I F P K E C V Y I H D P T G L N V L K K	324
1141	GGGCTGTGCCAGCTACTGCAACCAAAACCATCATGGAACAAGGCTGCTGCAAAGGTTTMTT	1200
325	G C A S Y C N Q T I M E Q G C C K G F F	344
1201	CGGGCCTGACTGCACGCAGTGTCTGGGGGCTTCTCCAACCCCTGCTATGGCAAAGGCAA	1260
345	G P D C T Q C P G G F S N P C Y G K G N	364
1261	TTGCACTGATGGGATCCAGGGCAATGGGGCCTGCCTCTGCTTCCCAGACTACAAGGGCAT	1320
365	C S D G I Q G N G A C L C F P D Y K G I	384
1321	CGCCTGCCACATCTGCTCGAACCCTAAACAAGCATGGAGAGCAATGCCAGGAAGACTGCGG	1380
385	A C H I C S N P N K H G E Q C Q E D C G	404
1381	CTGTGTCCATGGTCTCTGCGACAACCGCCAGGCAGTGGGGGGTGTGCCAGCAGGGCAC	1440
405	C V H G L C D N R P G S G G V C Q Q G T	424
1441	GTGTGCCCCTGGCTTCAGTGGCCGGTCTGCAACGAGTCCATGGGGGACTGTGGGCCCAC	1500
425	C A P G F S G R F C N E S M G D C G P T	444
1501	AGGGCTGGCCAGCACTGCCACCTGCATGCCCGCTGTGTAGCCAGGAGGGTGTGCCAG	1560
445	G L A Q H C H L H A R C V S Q E G V A R	464
1561	ATGTGCTGTCTTGATGGCTTTGAGGGTGATGGCTTCTCCTGCACACCTAGCAACCCCTG	1620
465	C R C L D G F E G D G F S C T P S N P C	484
1621	CTCCCAACCGGACCGTGGAGGCTGCTCAGAGAATGCTGAGTGTGCCCTGGGTCCCTGGG	1680
485	S H P D R G G C S E N A E C V P G S L G	504

Figure 1C

1681	CACCCACCACTGCACATGCCACAAAGGCTGGAGTGGGGATGGCCGCTCTGTGTGGCTAT	1740
505	T H H C T C H K G W S G d g x y c v a I	524
1741	TGACGAGTGTGAGCTGGACGTGAGAGGTGGCTGCCACACCGATGCCCTCTGCAGCTATGT	1800
525	D E C E L D V R G G C H T D A L C S Y V	544
1801	GGGCCCCGGGCAGAGCCGATGCACCTGCAAGCTGGGCTTTGCCGGGGATGGCTACCACTG	1860
545	G P G Q S R C T C K L G F A G D G Y Q C	564
1861	CAGCCCCATCGACCCCTGCCGGGCAGGCAATGGCGGCTGCCACGGCCTGGAGCTGGAGGC	1920
565	S P I D P C R A G N G G C H G L E L E A	584
1921	AAATGCCCACTTCTCCATCTTCTACCAATGGCTTAAGAGTGCCGGCATCACGCTTCCTGC	1980
585	N A H F S I F Y Q W L K S A G I T L P A	604
1981	CGACCGCCGAGTCACAGCCCTGGTGCCCTCCGAGGCTGCAGTCCGTCAGCTGAGCCCCGA	2040
605	D R R V T A L V P S E A A V R Q L S P E	624
2041	GGACCGAGCTTTCTGGCTGCAGCCAAGGACGCTGCCGAACCTGGTCAGGGCCCATTTTCT	2100
625	D R A F W L Q P R T L P N L V R A H F L	644
2101	CCAGGGTGCCCTCTTCGAGGAGGAGCTGGCCCCGGCTGGGTGGGCAGGAAGTGCCACCCCT	2160
645	Q G A L F E E E L A R L G G Q E V A T L	664
2161	GAACCCACCAACGCTGGGAGATTTCGCAACATTAGTGGGAGGGTCTGGGTGCAGAATGC	2220
665	N P T T R W E I R N I S G R V W V Q N A	684
2221	CAGCGTGGATGTGGCTGACCTCCTTGCCACCAACGGTGTCTACACATCCTCAGCCAGGT	2280
685	S V D V A D L L A T N G V L H I L S Q V	704
2281	CTTACTGCCCCCCCCGAGGGGATGTGCCCGGTGGGCAGGGGTGCTGCAGCAGCTGGACTT	2340
705	L L P P R G D V P G G Q G L L Q Q L D L	724
2341	GGTGCTGCCTTTCAGCCTCTTCCGGGAATTGCTGCAGCACCATGGGTGGTGCCCCAGAT	2400
725	V P A F S L F R E L L Q H H G L V P Q I	744
2401	TGAGGCTGCCACTGCCTACACCATCTTTGTGCCCCACCAACCGCTCCCTGGAGGCCCAGGG	2460
745	E A A T A Y T I F V P T N R S L E A Q G	764
2461	CAACAGCAGTCACCTGGACGCAGACACAGTGGGACCATGTGGTCTGGGGGAGGCCCT	2520
765	N S S H L D A D T V R H H V V L G E A L	784

Figure 1D

2521	CTCCATGGAAACCTGCGGAAGGGTGGACACCGCAACTCCCTCCTGGGCCCTGCCCACTG	2580
785	S M E T L R K G G H R N S L L G P A H W	804
2581	GATCGTCTTCTACAACCACAGTGGCCAGCCTGAGGTGAACCATGTGCCACTGGAAGGCCC	2640
805	I V F Y N H S G Q P E V N H V P L E G P	824
2641	CATGCTGGAGGCCCTGGCCGCTCGCTGATTGGTCTGTGCGGGGTCCTGACGGTGGGCTC	2700
825	M L E A P G R S L I G L S G V L T V G S	844
2701	AAGTCGCTGCCTGCATAGCCACGCTGAGGCCCTGCGGGAGAAATGTGTAAACTGCACCAG	2760
845	S R C L H S H A E A L R E K C V N C T R	864
2761	GAGATTCCGCTGCACTCAGGGCTTCCAGCTGCAGGACACACCCAGGAAGAGCTGTGTCTA	2820
865	R F R C T Q G F Q L Q D T P R K S C V Y	884
2821	CCGATCTGGCTTCTCCTTCTCCCGGGCTGCTCTTACACATGTGCCAAGAAGATCCAGGT	2880
885	R S G F S F S R G C S Y T C A K K I Q V	904
2881	GCCGGACTGCTGCCCTGGTTTCTTTGGCACGCTGTGTGAGCCATGCCCAGGGGGTCTAGG	2940
905	P D C C P G F F G T L C E P C P G G L G	924
2941	GGGGGTGTGCTCAGGCCATGGGCAGTGCCAGGACAGGTTCTTGGGCAGCGGGGAGTGCCA	3000
925	G V C S G H G Q C Q D R F L G S G E <u>C</u> <u>H</u>	944
3001	CTGCCACGAGGGCTTCCATGGAACGGCCTGTGAGGTGTGTGAGCTGGGCCGCTACGGGCC	3060
945	<u>C</u> <u>H</u> <u>E</u> <u>G</u> <u>F</u> <u>H</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>E</u> <u>V</u> <u>C</u> <u>E</u> <u>L</u> <u>G</u> <u>R</u> <u>Y</u> <u>G</u> <u>P</u>	964
3061	CAACTGCACCGGAGTGTGTGACTGTGCCCATGGGCTGTGCCAGGAGGGGCTGCAAGGGGA	3120
965	N C T G V C D C A H G L C Q E G L Q G D	984
3121	CGGAAGCTGTGTCTGTAACGTGGGCTGGCAGGGCCTCCGCTGTGACCAGAAAATCACCAG	3180
985	G S <u>C</u> <u>V</u> <u>C</u> <u>N</u> <u>V</u> <u>G</u> <u>W</u> <u>Q</u> <u>G</u> <u>L</u> <u>R</u> <u>C</u> <u>D</u> <u>Q</u> <u>K</u> <u>I</u> <u>T</u> <u>S</u>	1004
3181	CCCTCAGTCCCCTAGGAAGTGCGACCCCAATGCCAACTGCGTGCAGGACTCGGCCGGAGC	3240
1005	P Q C P R K C D P N A N C V Q D S A G A	1024
3241	CTCCACCTGCGCCTGTGCTGCCGGATACTCCGGCAATGGCATCTTCTGTTCAGAGGTGGA	3300
1025	S T <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>Y</u> <u>S</u> <u>G</u> <u>N</u> <u>G</u> <u>I</u> <u>F</u> <u>C</u> <u>S</u> <u>E</u> <u>V</u> <u>D</u>	1044
3301	CCCCTGCGCCACGGCCATGGGGGCTGTCCCCTCATGCCAACTGTACCAAGGTGGCACC	3360
1045	P C A H G H G G C S P H A N C T K V A P	1064

Figure 1E

3361	TGGGCAGCGGACATGCACCTGCCAGGATGGCTACATGGGCGACGGGGAGCTGTGCCAGGA	3420
1065	G Q R T C T C Q D G Y M G D G E L C Q E	1084
3421	AATTAACAGCTGTCTCATCCACCACGGGGGCTGCCACATTACGCCGAGTGCATCCCCAC	3480
1085	I N S C L I H H G G C H I H A E C I P T	1104
3481	TGGCCCCCAGCAGGTCTCCTGCAGCTGCCGTGAGGGTTACAGCGGGGATGGCATCCGGAC	3540
1105	G P Q Q V S C S C R E G Y S G D G I R T	1124
3541	CTGCGAGCTCCTGGACCCCTGCTCTAAGAACAATGGAGGATGCAGCCCATATGCCACCTG	3600
1125	C E L L D P C S K N N G G C S P Y A T C	1144
3601	CAAAAGCACAGGGGATGGCCAGAGGACATGTACCTGCGACACAGCCACACCGTGGGGGA	3660
1145	K S T G D G Q R T C T C D T A H T V G D	1164
3661	CGGCCTCACCTGCCGTGCCCGAGTCGGCCTGGAGCTCCTGAGGGATAAGCATGCCTCATT	3720
1165	G L T C R A R V G L E L L R D K H A S F	1184
3721	CTTCAGCCTCCGCCTCCTGGAATATAAGGAGCTCAAGGGCGATGGGCCTTTCACCATCTT	3780
1185	F S L R L L E Y K E L K G D G P F T I F	1204
3781	CGTGCCCGCAGCAGATCTAATGAGCAACCTGTGCGAGGATGAGCTGGCCCGGATTCTGTGC	3840
1205	V P H A D L M S N L S Q D E L A R I R A	1224
3841	GCATCGCCAGCTGGTGTTCGCTACCACGTGGTTGGCTGTGCGCGGCTGCGGAGCGAGGA	3900
1225	H R Q L V F R Y H V V G C R R L R S E D	1244
3901	CCTGCTGGAGCAGGGGTACGCCACGGCCCTCTCAGGGCACCCACTGCGCTTCAGCGAGAG	3960
1245	L L E Q G Y A T A L S G H P L R F S E R	1264
3961	GGAGGGCAGCATATACCTCAATGACTTCGCGCGCGTGGTGAGCAGCGACCATGAGGCCGT	4020
1265	E G S I Y L N D F A R V V S S D H E A V	1284
4021	GAACGGCATCCTGCACTTCATTGACCGTGTCTGCTGCCCCCGAGGCGCTGCACTGGGA	4080
1285	N G I L H F I D R V L L P P E A L H W E	1304
4081	GCCTGATGATGCTCCCATCCCGAGGAGAAATGTCACCGCCGCGCCAGGGCTTCGGTTA	4140
1305	P D D A P I P R R N V T A A A Q G F G Y	1324
4141	CAAGATCTTCAGCGGCCTCCTGAAGGTGGCCGGCCTCCTGCCCCCTGCTTCGAGAGGCATC	4200
1325	K I F S G L L K V A G L L P L L R E A S	1344

Figure 1F

4201	CCATAGGCCCTTCACAATGCTGTGGCCACAGACGCCGCTTTCGAGCTCTGCCTCCGGA	4260
1345	H R P F T M L W P T D A A F R A L P P D	1364
4261	TCGCCAGGCCTGGCTGTACCATGAGGACCACCGTGACAAGCTAGCAGCCATTCTGCGGGG	4320
1365	R Q A W L Y H E D H R D K L A A I L R G	1384
4321	CCACATGATTGCAATGTCGAGGCCTTGGCATCTGACCTGCCCAACCTGGGCCCCACTTCG	4380
1385	H M I R N V E A L A S D L P N L G P L R	1404
4381	AACCATGCATGGGACCCCATCTCTTTCTCCTGCAGCCGAACGCGGCCCGGTGAGCTCAT	4440
1405	T M H G T P I S F S C S R T R P G E L M	1424
4441	GGTGGGTGAGGATGATGCTCGCATTTGTGCAGCGGCACCTTGCCCTTTGAGGGTGGCCTGGC	4500
1425	V G E D D A R I V Q R H L P F E G G L A	1444
4501	CTATGGCATCGACCAGCTGCTGGAGCCACCTGGCCTTGGTGCTCGCTGTGACCACTTTGA	4560
1445	Y G I D Q L L E P P G L G A R C D H F E	1464
4561	GACCCGGCCCTGCGACTGAACACCTGCAGCATCTGTGGGCTGGAGCCACCTGTCCTGA	4620
1465	T R P L R L N T C S I C G L E P P C P E	1484
4621	GGGGTCACAGGAGCAGGGCAGCCCTGAGGCCTGCTGGCGCTTCTACCCGAAGTTCTGGAC	4680
1485	G S Q E Q G S P E A C W R F Y P K F W T	1504
4681	GTCCCTCCGCTGCACTCTTTGGGATTACGCAGCGTCTGGGTCCACCCAGCCTTTGGGG	4740
1505	S P P L H S L G L R S V W V H P S L W G	1524
4741	TAGGCCCAAGGCCTGGGCAGGGGCTGCCACCGCAATTGTGTCAACCACCACTGGAAGCC	4800
1525	R P Q G L G R G C H R N C V T T T W K P	1544
4801	CAGCTGCTGCCCTGGTCACTATGGCAGTGAGTGCCAAGCTTGCCCTGGCGGCCCCAGCAG	4860
1545	S C C P G H Y G S E C Q A C P G G P S S	1564
4861	CCCTTGTAAGTGACCGTGGCGTGTGCATGGACGGCATGAGTGGCAGTGGGCAGTGTCTGTG	4920
1565	P C S D R G V C M D G M S G S G Q C <u>L</u> <u>C</u>	1584
4921	CCGTTTCAGGTTTTCCTGGGACAGCCTGTGAACCTCTGTGCTCCTGGTGCCTTTGGGCCCCA	4980
1585	R <u>S</u> <u>G</u> <u>F</u> <u>A</u> <u>G</u> <u>T</u> <u>A</u> <u>C</u> <u>E</u> <u>L</u> <u>C</u> A P G A F G P H	1604
4981	TTGTCAAGCCTGCCGCTGCACTGTGCATGGCCGCTGTGATGAGGGCCTTGGGGGCTCTGG	5040
1605	C Q A C R C T V H G R C D E G L G G S G	1624

Figure 1G

5041	CTCCTGCTTCTGTGATGAAGGCTGGACTGGGCCACGCTGTGAGGTGCAACTGGAGCTGCA	5100
1625	S C F C D E G W T G P R C E V Q L E L Q	1644
5101	GCCTGTGTGTACCCACCCCTGTGCACCCGAGGCTGTGTGCCGTGCAGGCAACAGCTGTGA	5160
1645	P V C T P P C A P E A V C R A G N S C E	1664
5161	GTGCAGCCTGGGCTATGAAGGGGATGGCCGTGTGTGTACAGTGGCAGACCTGTGCCAGGA	5220
1665	C S L G Y E G D G R V C T V A D L C Q D	1684
5221	CGGGCATGGTGGCTGCAGTGAGCACGCCAACTGTAGCCAGGTAGGAACAATGGTCACTTG	5280
1685	G H G G C S E H A N C S Q V G T M V T C	1704
5281	TACCTGCCTGCCCGACTACGAGGGTGTGGCTGGAGCTGCCGGGCCGCAACCCTGCAC	5340
1705	T C L P D Y E G D G W S C R A R N P C T	1724
5341	AGATGGCCACCGCGGGGGTGCAGCGAGCACGCCAACTGCTTGAGCACCGGCCTGAACAC	5400
1725	D G H R G G C S E H A N C L S T G L N T	1744
5401	ACGGCGCTGTGAGTGCCACGCAGGCTACGTAGGCGATGGACTGCAGTGTCTGGAGGAGTC	5460
1745	R R C E C H A G Y V G D G L Q C L E E S	1764
5461	GGAACCACCTGTGGACCGCTGCTTGGGCCAGCCACCGCCCTGCCACTCAGATGCCATGTG	5520
1765	E P P V D R C L G Q P P P C H S D A M C	1784
5521	CACTGACCTGCACCTTCCAGGAGAAACGGGCTGGCGTTTTTCCACCTCCAGGCCACCAGCGG	5580
1785	T D L H F Q <u>E K R A G V F H L Q A T S G</u>	1804

5581	CCCTTATGGTCTGAACTTTTCGGAGGCTGAGGCGGCATGCGAAGCACAGGAGCCGCTCCT	5640
1805	<u>P Y G L N F S E A E A A C E A Q G A V L</u>	1824

5641	TGCTTCATTCCTCAGCTCTCTGCTGCCAGCAGCTGGGCTTCCACCTGTGCCTCATGGG	5700
1825	<u>A S F P Q L S A A Q Q L G F H L C L M G</u>	1844

5701	CTGGCTGGCCAATGGCTCCACTGCCACCCTGTGGTTTTTCCCTGTGGCGGACTGTGGCAA	5760
1845	<u>W L A N G S T A H P V V F P V A D C G N</u>	1864
5761	TGGTCGGGTGGGCATAGTCAGCCTGGGTGCCCGCAAGAACCTCTCAGAACGCTGGGATGC	5820
1865	<u>G R V G I V S L G A R K N L S E R W D A</u>	1884

Figure 1H

5821	CTACTGCTTCCGTGTGCAAGATGTGGCCTGCCGATGCCGAAATGGCTTCGTGGGTGACGG	5880
1885	<u>Y C F R V Q D V A C R C R N G E V G D G</u>	1904
5881	GATCAGCACGTGCAATGGGAAGCTGCTGGATGTGCTGGCTGCCACTGCCAACTTCTCCAC	5940
1905	I S T C N G K L L D V L A A T A N F S T	1924
5941	CTTCTATGGGATGCTATTGGGCTATGCCAATGCCACCCAGCGGGGTCTCGACTTCCTGGA	6000
1925	F Y G M L L G Y A N A T Q R G L D F L D	1944
6001	CTTCCTGGATGATGAGCTCACGTATAAGACACTCTTCGTCCTGTCAATGAAGGCTTTGT	6060
1945	F L D D E L T Y K T L F V P V N E G F V	1964
6061	GGACAACATGACGCTGAGTGGCCCAAACCTGGAGCTGCATGCCTCCAACGCCACCCTCCT	6120
1965	D N M T L S G P K L E L H A S N A T L L	1984
6121	AAGTGCCAACGCCAGCCAGGGGAAGTTGCTTCCGGCCCACTCAGGCCTCAGCCTCATCAT	6180
1985	S A N A S Q G K L L P A H S G L S L I I	2004
6181	CAGTGACGCAGGCCCTGACAACAGTTCTGGGCCCTGTGGCCCCAGGGACAGTTGTGGT	6240
2005	S D A G P D N S S W A P V A P G T V V V	2024
6241	TAGCCGTATCATTTGTGTGGGACATCATGGCCTTCAATGGCATCATCCATGCTCTGGCCAG	6300
2025	S R I I V W D I M A F N G I I H A L A S	2044
6301	CCCCCTCCTGGCACCCCCACAGCCCCAGGCAGTGCTGGCGCNTGAAGCCCCACCTGTGGC	6360
2045	P L L A P P Q P Q A V L A X E A P P V A	2064
6361	GGCAGGCGTGGGGGCTGTGCTTGGCGCTGGAGCACTGCTTGGCTTGGTGGCCGGAGCTCT	6420
2065	A G V G A V L A A G A L L G L V A G A L	2084
6421	CTACCTCCGTGCCCGAGGCAAGCCCATGGGCTTTGGCTTCTCTGCCTTCCAGGCGGAAGA	6480
2085	Y L R A R G K P M G F G F S A F Q A E D	2104
6481	TGATGCTGATGACGANTTCTCACCGTGGCAAGAAGGGACCAACCCACANTTGGTNTNTGT	6540
2105	D A D D X F S P W Q E G T N P T L V X V	2124
6541	CCCCAACCTGTCTTTGGCAGCGACACCTTTTGTGAACCTTCGATGACTCACTGCTGGA	6600
2125	P N P V F G S D T F C E P F D D S L L E	2144
6601	GGAGGACTTCCCTGACACCCAGAGGATCCTCACAGTCAAGTGACGAGGCTGGGGCTGAAA	6660
2145	E D F P D T Q R I L T V K *	2158

6661 GCAGAAGCATGCACAGGGAGGAGACCANTTTTATTGCTTGTCTGGGTGGATGGGGCAGGA 6720

6721 GGGNCTGAGGGCCTGTCCCAGACAATANNNGTNCCTCGAG 6761

Figure 2A

1	GAGCACGCCAACTGTAGCCAGGTAGGAACAATGGTCACTTGTACCTGCCCTGCCCCGACTAC	60
1	M V T C T C L P D Y	10
61	GAGGGTGATGGCTGGAGCTGCCGGGCCCCGAACCCCTGCACAGATGGCCACCGCGGGGGC	120
11	E G D G W S C R A R N P C T D G H R G G	30
121	TGCAGCGAGCACGCCAACTGCTTGAGCACCGGCCTGAACACACGGCGCTGTGAGTGCCAC	180
31	C S E H A N C L S T G L N T R R C E C H	50
181	GCAGGCTACGTAGGCGATGGACTGCAGTGTCTGGAGGAGTCGGAACCACCTGTGGACCGC	240
51	A G Y V G D G L Q C L E E S E P P V D R	70
241	TGCTTGGGGCAGCCACCGCCCTGCCACTCAGATGCCATGTGCACTGACCTGCACTTCCAG	300
71	C L G Q P P P C H S D A M C T D L H F Q	90
301	GAGAAACGGGCTGGCGTTTTCACCTCCAGGCCACCAGCGGCCCTTATGGTCTGAAC'TTT	360
91	<u>E K R A G V F H L Q A T S G P Y G L N F</u>	110

361	TCGGAGGCTGAGGCGGCATGCCAAGCACAGGGAGCCGTCCTTGCTTCATTCCCTCAGCTC	420
111	<u>S E A E A A C E A Q G A V L A S F P Q L</u>	130

421	TCTGCTGCCCAGCAGCTGGGCTTCCACCTGTGCCTCATGGGCTGGCTGGCCAATGGCTCC	480
131	<u>S A A Q Q L G F H L C L M G W L A N G S</u>	150

481	ACTGCCACCCCTGTGGTTTTCCTGTGGCGGACTGTGGCAATGGTCGGGTGGGCATAGTC	540
151	<u>T A H P V V F P V A D C G N G R V G I V</u>	170
541	AGCCTGGGTGCCCCGAAGAACCCTCTCAGAACGCTGGGATGCCTACTGCTTCCGTGTGCAA	600
171	<u>S L G A R K N L S E R W D A Y C F R V Q</u>	190
601	GATGTGGCCTGCCGATGCCGAATGGCTTCGTGGGTGACGGGATCAGCACGTGCAATGGG	660
191	<u>D V A C R C R N G F Y G D G I S T C N G</u>	210
661	AAGTGCTGGATGTGCTGGCTGCCACTGCCAACTTCTCCACCTTCTATGGGATGCTATTG	720
211	K L L D V L A A T A N F S T F Y G M L L	230
721	GGCTATGCCAATGCCACCCAGCGGGTCTCGACTTCCTGGACTTCCTGGATGATGAGCTC	780
231	G Y A N A T Q R G L D F L D F L D D E L	250
781	ACGTATAAGACACTCTTCGTCCCTGTCAATGAAGGCTTTGTGGACAACATGACGCTGAGT	840

251 T Y K T L F V P V N E G F V D N M T L S 270

Figure 2B

841 GGCCCAAACCTGGAGCTGCATGCCTCCAACGCCACCCTCCTAAGTGCCAACGCCAGCCAG 900
271 G P N L E L H A S N A T L L S A N A S Q 290

901 GGGAAGTTGCTTCCGGCCCACTCAGGCCTCAGCCTCATCATCAGTGACGCAGGCCCTGAC 960
291 G K L L P A H S G L S L I I S D A G P D 310

961 AACAGTTCTGGGCCCCCTGTGGCCCCAGGGACAGTTGTGGTTAGCCGTATCATTGTGTGG 1020
311 N S S W A P V A P G T V V V S R I I V W 330

1021 GACATCATGGCCTTCAATGGCATCATCCATGCTCTGGCCAGCCCCCTCCTGGCACCCCCA 1080
331 D I M A F N G I I H A L A S P L L A P P 350

1081 CAGCCCCAGGCAGTGCTGGCGCNTGAAGCCCCACCTGTGGCGGCAGGCGTGGGGGCTGTG 1140
351 Q P Q A V L A X E A P P V A A G V G A V 370

1141 CTTGCCGCTGGAGCACTGCTTGGCTTGGTGGCCGGAGCTCTCTACCTCCGTGCCCGAGGC 1200
371 L A A G A L L G L V A G A L Y L R A R G 390

1201 AAGCCCATGGGCTTGTGGCTTCTCTGCCTTCCAGGCGGAAGATGATGCTGATGACGANITC 1260
391 K P M G F G F S A F Q A E D D A D D X F 410

1261 TCACCGTGGCAAGAAGGGACCAACCCACNITGGTNTMTGTCCCAACCCTGTCTTTGGC 1320
411 S P W Q E G T N P T L V X V P N P V F G 430

1321 AGCGACACCTTTTGTGAACCCTTCGATGACTCACTGCTGGAGGAGGACTTCCCTGACACC 1380
431 S D T F C E P F D D S L L E E D F P D T 450

1381 CAGAGGATCCTCACAGTCAAGTGACGAGGCTGGGGCTGAAAGCAGAAGCATGCACAGGGA 1440
451 Q R I L T V K * 458

1441 GGAGACCANITTTATTGCTTGTCTGGGTGGATGGGGCAGGAGGGNCTGAGGGCCTGTCCC 1500

1501 AGACAATANNGTNCCTCGAG 1522

Figure 3A

```

1  GCCCACGCGTCCGACCGGGACAGCTCGCGGCCCCCNAGAGCTCTAGCCGTNGAGGAGCTG  60
61  CCTGGGGACGTTTGCCCTGGGGCCCCAGCCTGGCCCGGGTCACCCTGGCATGAGGAGATG  120
1  M 1
121  GGCCTGTGTGCTCCTGGTCCCATTTGCTCCTGCTGCCCCGGCTCCTACGGACTGCCCTTCTAC  180
2  G L L L L V P L L L L P G S Y G L P F Y  21
181  TACGGCTTCTACTACTCCAACAGCGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAA  240
22  Y G F Y Y S N S A N D Q N L G N G H G K  41
241  GACCTACNTAATGGAGTGAAGCTGGTGGTGGAGACACCCGAGGAGACCCTGTTACCTAC  300
42  D L X N G V K L V V E T P E E T L F T Y  61
301  CAAGGGGCCAGTGTGATCCTGCCCCTGCGCTACCGCTACGAGCCGGCCCTGGTCTCCCCCG  360
62  Q G A S V I L P C R Y R Y E P A L V S P  81
361  CGGCGTGTGCGTGTCAAATGGTGAAGCTGTGCGGAGAACGGGGCCCCAGAGAAGGACGTG  420
82  R R V R V K W W K L S E N G A P E K D V 101
421  CTGGTGGCCATCGGGCTGAGGCACCGCTCCTTTGGGGACTACCAAGGCCGCGTGCACCTG  480
102  L V A I G L R H R S F G D Y Q G R V H L 121
481  CGGCAGGACAAAGAGCATGACGTCTCGNTGGAGATCCAGGNTCTGCGGCTGGAGGACTAT  540
122  R Q D K E H D V S X E I Q X L R L E D Y 141
541  GGGCGTTACCGCTGTGAGGTCATNGACGGGCTGGAGGATGAAAGCGGTCTGGTGGAGCTG  600
142  G R Y R C E V X D G L E D E S G L V E L 161
601  GAGCTGCGGGGTGTGGTCTTTCTTACCACTCCCCCAACGGGCGCTACCACTTCAACTTC  660
162  E L R G V V F P Y Q S P N G R Y Q F N F 181
*****
661  CACGAGGGCCAGCAGGTCTGTGACAGCAGGCTGCGGTGGTGGCCTCCTTTGAGCAGCTC  720
182  H E G Q Q V C A E Q A A V V A S F E Q L 201
*****
721  TTCCGGGCTGGGAGGAGGGCTGGACTGGTGAACGCGGGCTGGCTGCAGGATGCCACG  780
202  F R A W E E G L D W C N A G W L Q D A T 221
*****
781  GTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCCCGGACCTGGCACCTGGC  840
222  V Q Y P I M L P R Q P C G G P D L A P G 241

```

Figure 3B

841 GTGCGAAGCTACGGCCCCCGCCACCGCGCCTGCACCGCTATGATGTATTCTGCTTCGCT 900
242 V R S Y G P R H R R L H R Y D V F C F A 261

901 ACTGCCCTCARGGGGCGGGTGTACTACCTGGANCAACCTGAGAANCTGACNCTGACANAA 960
262 T A L X G R V Y Y L X H P E X L T L T X 281

961 GCAAGGGAAGCCTGCCAAGAAAAAT 985
282 A R E A C Q E K 289

Figure 4A

1 GGAATCACATGCACAGTTGTGGATTtTGTGCAAACAGGACAACGGGGGCTGTGCAAAGGTG 60

61 GCCAGATGCTCCCAGAAGGGCACGAAGGTCTCCTGCAGCTGCCAGAAGGGATACAAAGGG 120

121 GACGGGCACAGCTGCACAGAGATAGACCCCTGTGCAGACGGCCTTAACGGAGGGTGTAC 180

181 GAGCACGCCACCTGTAAGATGACAGGCCCGGGCAAGCACAAGTGTGAGTGTAAAAGTCAC 240
1 M T G P G K H K C E C K S H 14

241 TATGTCGGAGATGGGCTGAACTGTGAGCCGGAGCAGCTGCCCATTGACCGCTGCTTACAG 300
15 Y V G D G L N C E P E Q L P I D R C L Q 34

301 GACAATGGGCAGTGCCATGCAGACGCCAAATGTGTGCGACCTCCACTTCCAGGATACCACT 360
35 D N G Q C H A D A K C V D L H F Q D T T 54

361 GTTGGGGTGTTCATCTACGCTCCCCACTGGGCCAGTATAAGCTGACCTTTGACAAAGCC 420
55 V G V F H L R S P L G Q Y K L T F D K A 74

421 AGAGAGGCCTGTGCCAACGAAGCTGCGACCATGGCAACCTACAACCAGCTCTCCTATNNC 480
75 R E A C A N E A A T M A T Y N Q L S Y X 94

481 CAGAAGGCCAAGTACCACCTGTGCTCAGCAGGCTGGCTGGAGACCGGGCGGGTTGCCTAC 540
95 Q K A K Y H L C S A G W L E T G R V A Y 114

541 CCCACAGCCTTCGCTCCAGAACTGTGGCTCTGGTGTGGTGGGATAGTGGACTATGGA 600
115 P T A F A S Q N C G S G V V G I V D Y G 134

601 CCTAGACCCAACAAGAGTGAAATGTGGGATGTCTTCTGCTATCGGATGAAAGATGTGAAC 660
135 P R P N K S E M W D V F C Y R M K D V N 154

661 TGCACCTNCAAGGTGGGCTATGTGGGAGATGGCTTCTCATaCAGTGGGAACCTGCTGCAG 720
155 C T X K V G Y V G D G F S Y S G N L L Q 174

721 GTCCTGATGTCCTTCCCCTCACTCACAACTTCTGACGGAAGTGTGGcCTATTCCAaC 780
175 V L M S F P S L T N F L T E V L A Y S N 194

781 AGcTCAGcTCGAGGCCGtGCATTtTCTAGAACACCTGACTGACCTGTCCATCCGcGGCACC 840
195 S S A R G R A F L E H L T D L S I R G T 214

Figure 4B

841	CTCTtTGTNcCACAGAACAGTGGgCTGGGGGAGAATGAGACCTTGTCTGGGcGGGACATC	900
215	<u>L</u> F V P Q N S G L G E N E T L S G R D I	234
901	GAGCACCACTCGCCAATGTCAGcATGTTTtTCTACAATGACctTgtCAATGGCACCACC	960
235	E H H L A N V S M F F Y N D L V N G T T	254
961	CTGCAAACGAGGcTGGGAAGCAAgCTGCTCATCACTGaCAGaCAGGACCCACTCCAcCCG	1020
255	L Q T R L G S K L L I T D R Q D P L H P	274
1021	AcGGAGACCAGGTgTGTtTGATGGAAGaGaCActCTGgAGtGGGACATCTgTGcCTCCAaT	1080
275	T E T R C V D G R D T L E W D I C A S N	294
1081	GGgatCACaCATGTCAttTCCAGGyCTTTAAaAGCACCCCTgCCCCCGtGAcCttGNCC	1140
295	G I T H V I S R X L K A P P A P V T L X	314
1141	CACACTGgNTtGGGAgNAGGGATCTTCTNTGNCATCATCCTGGTGACTGGGGCTGTtGCC	1200
315	H T G L G X G I F X X I I L V T G A V A	334
1201	TTGGCTGCTTACTCCTACTTTTCGGATAAACC GGAAAACAATCGGCTTCCANCATTTTGA	1259
335	L A A Y S Y F R I N R K T I G F X H F	353

Figure 5A

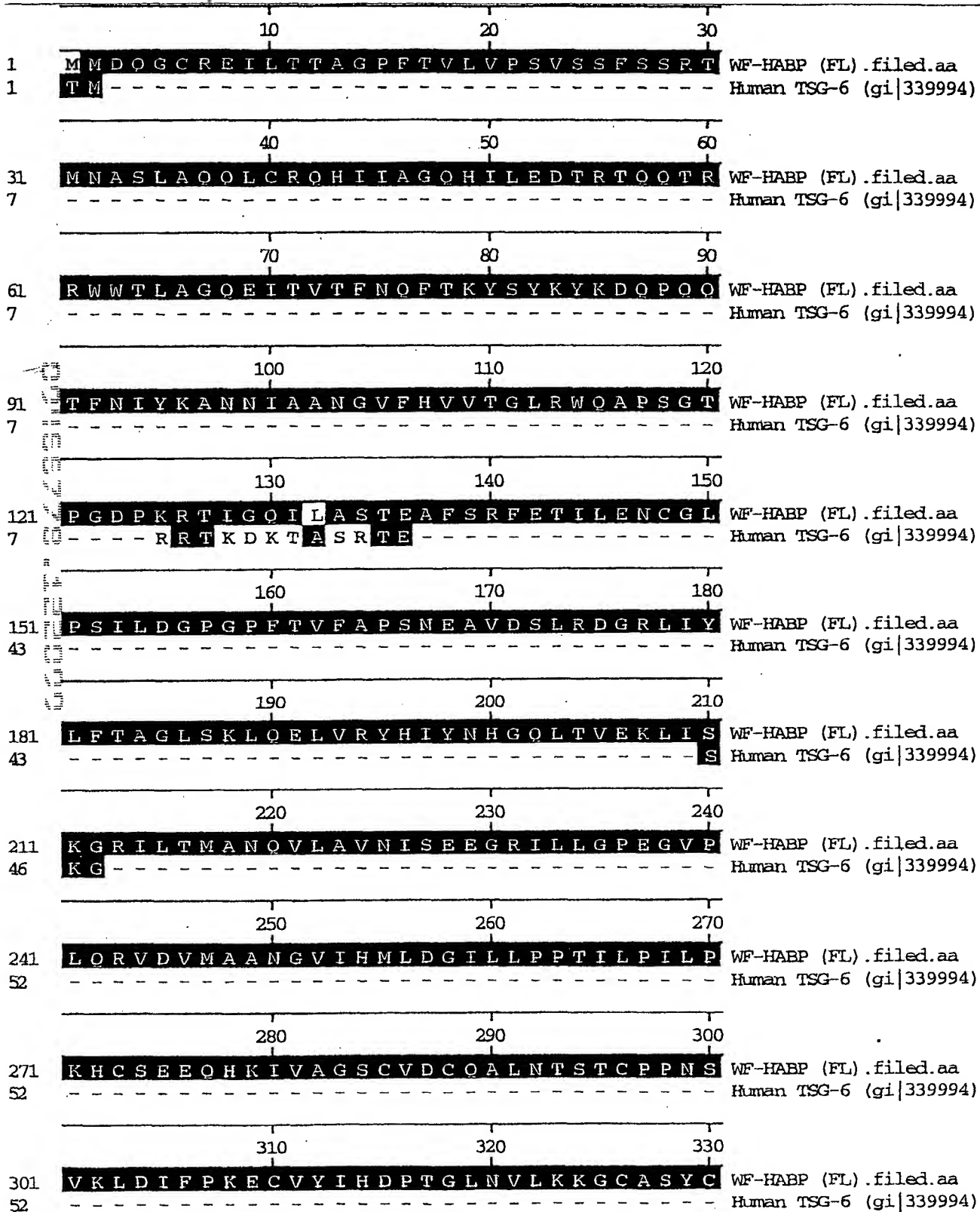


Figure 5B

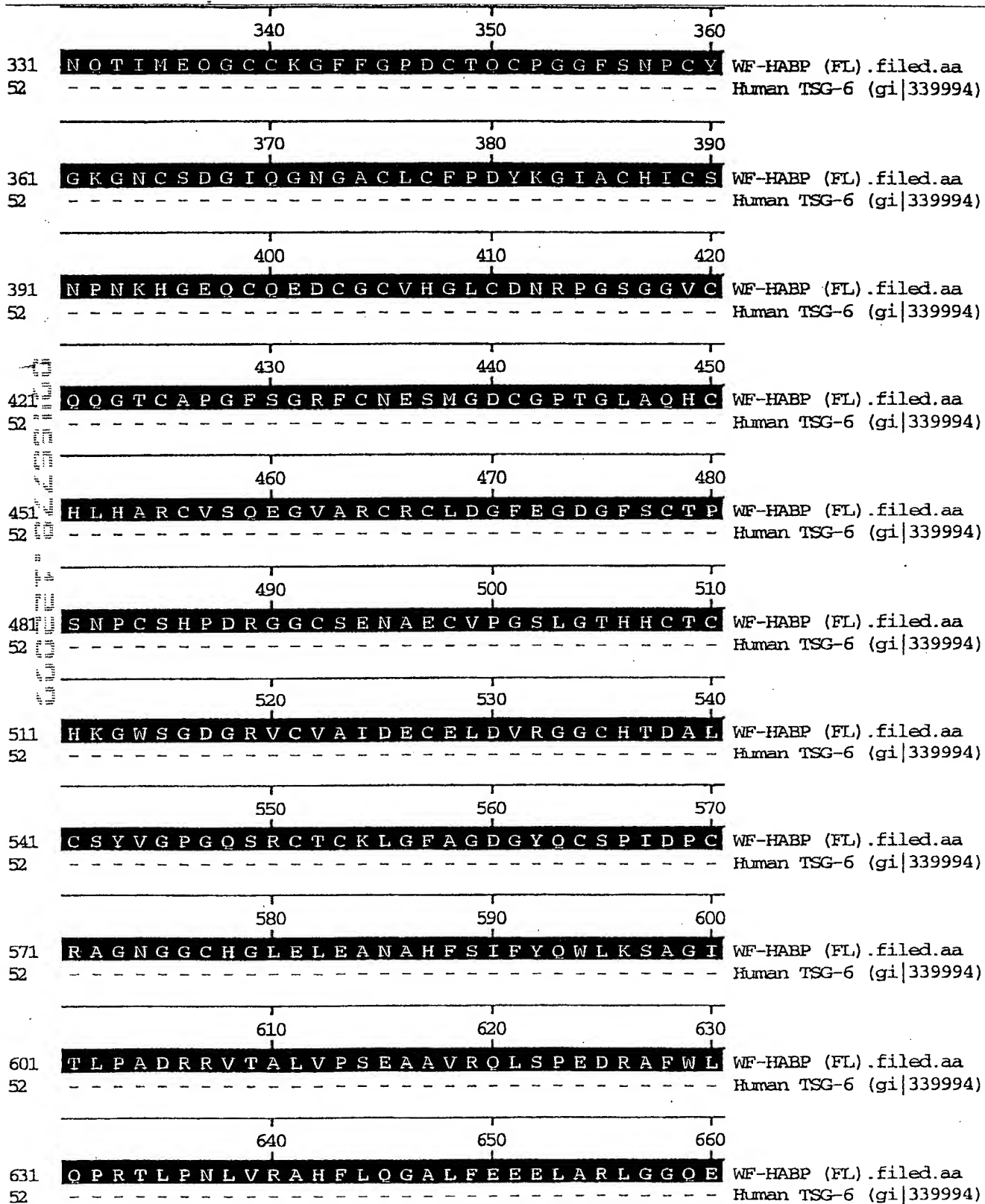


Figure 5C

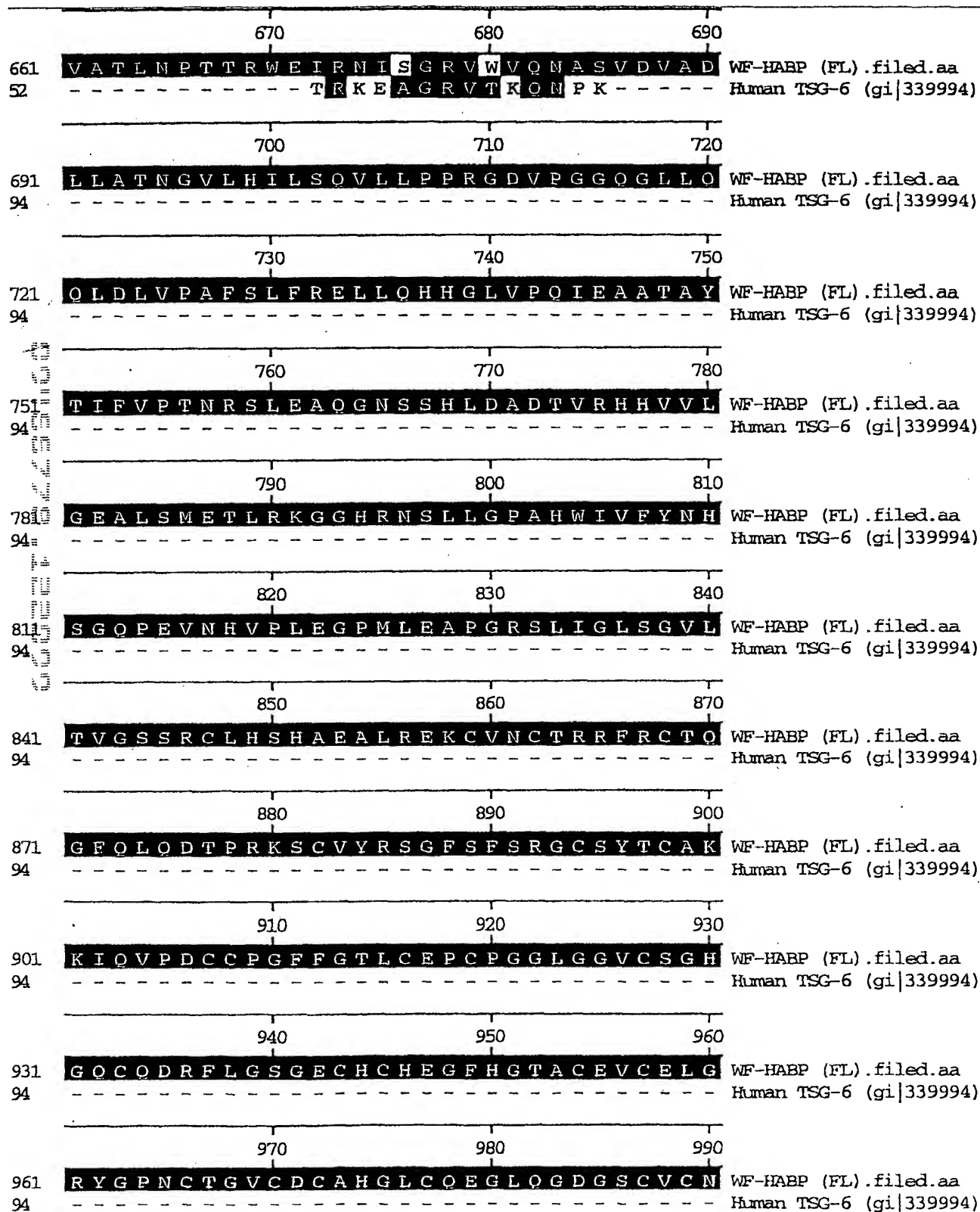


Figure 5D

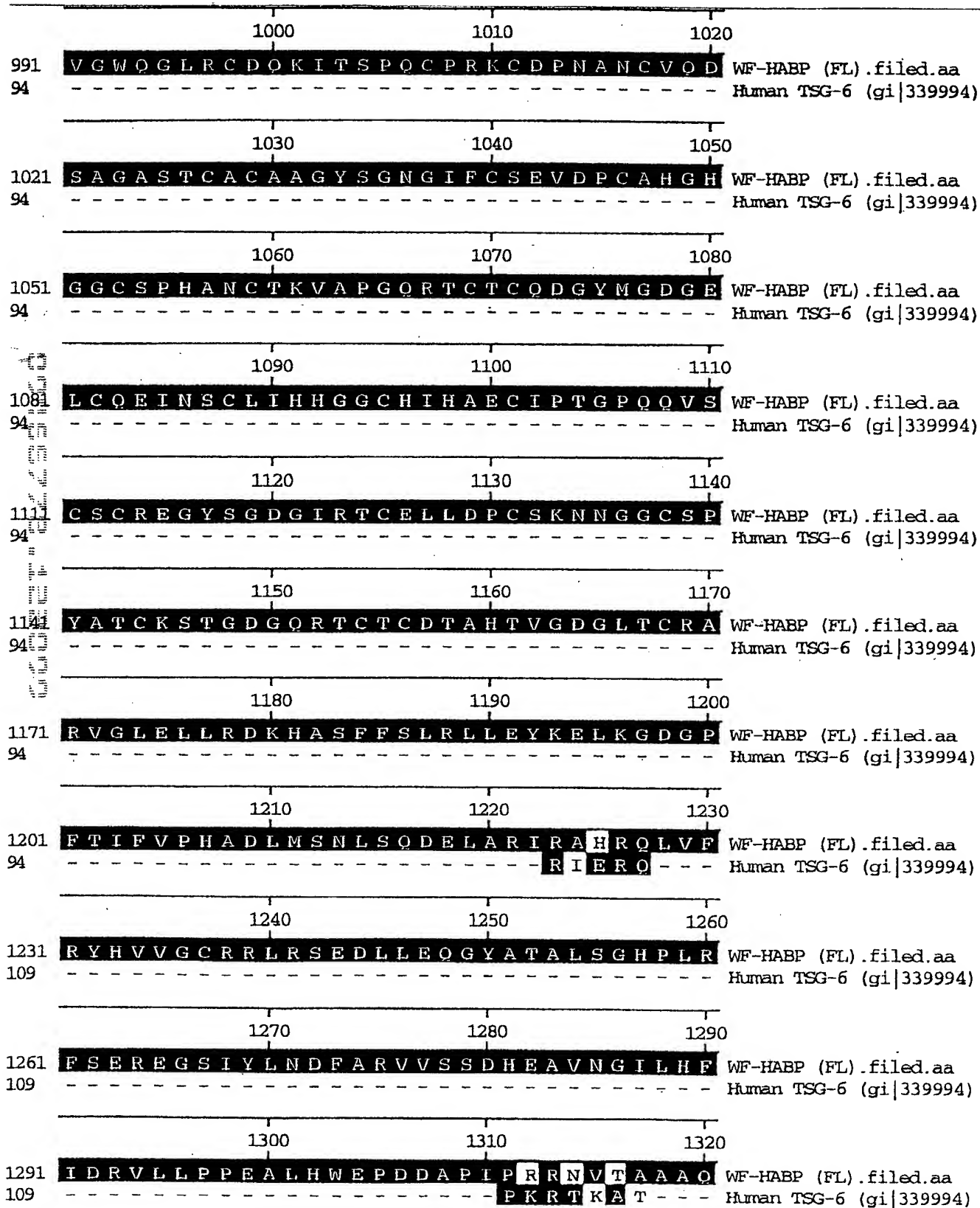


Figure 5E

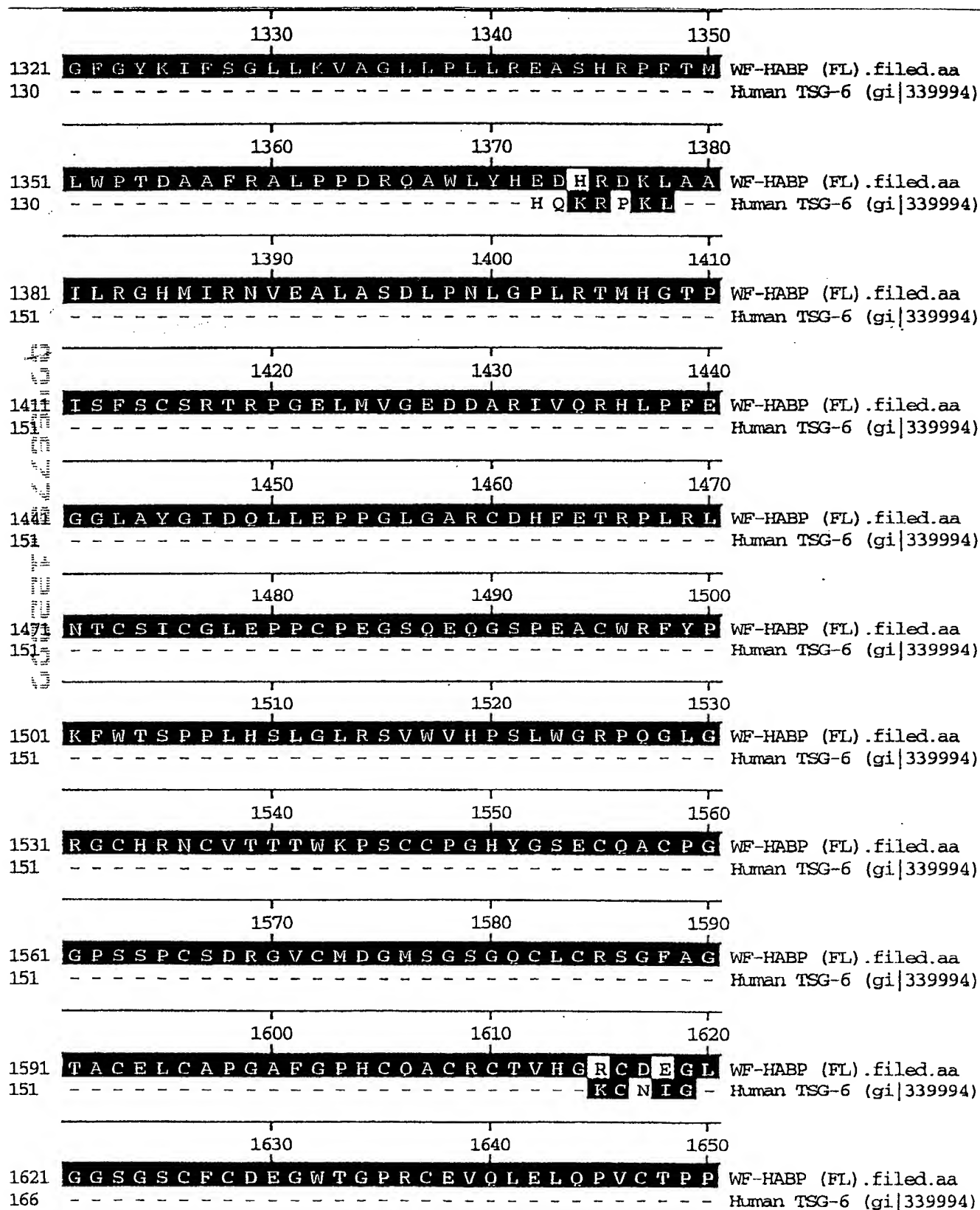


Figure 5F

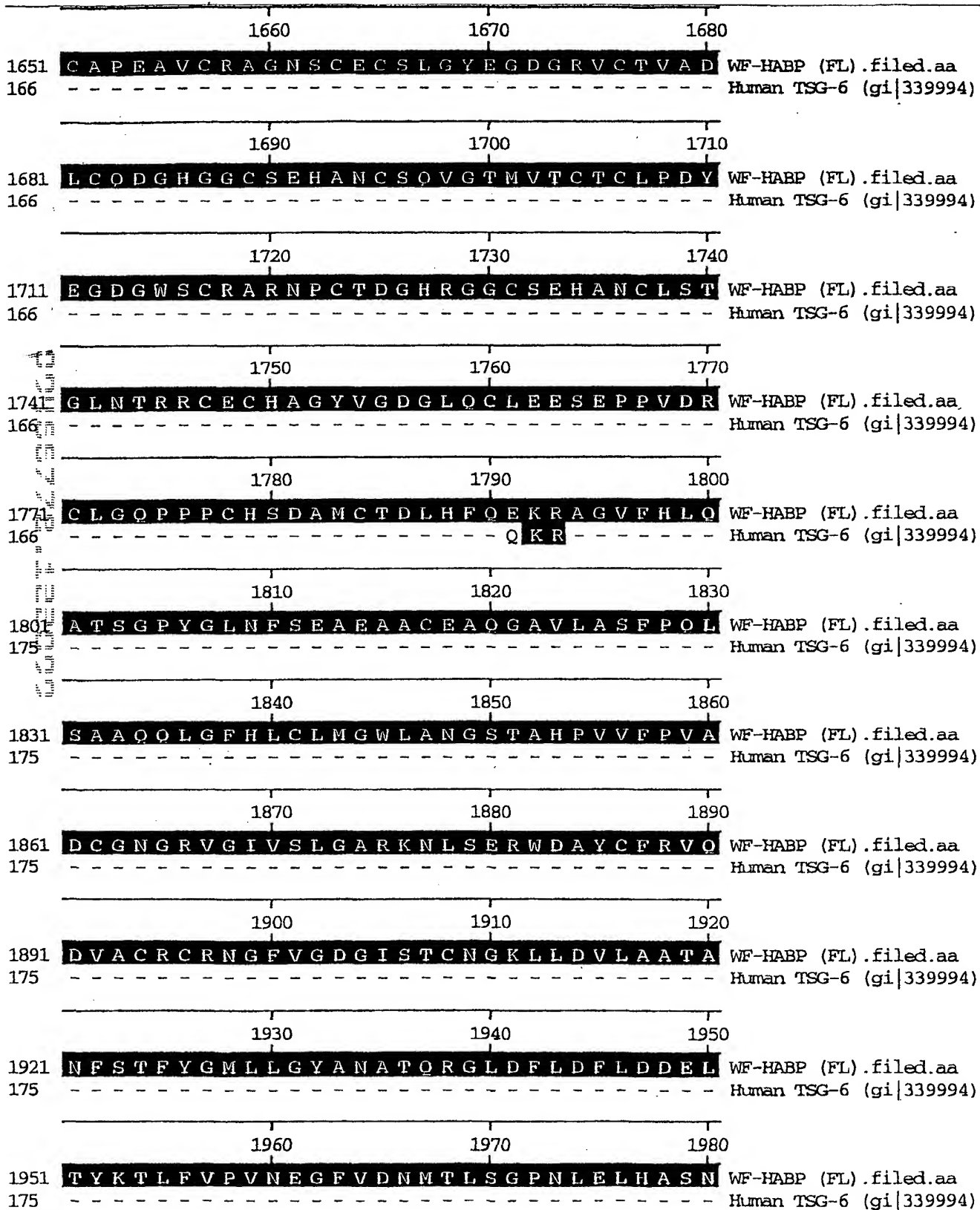
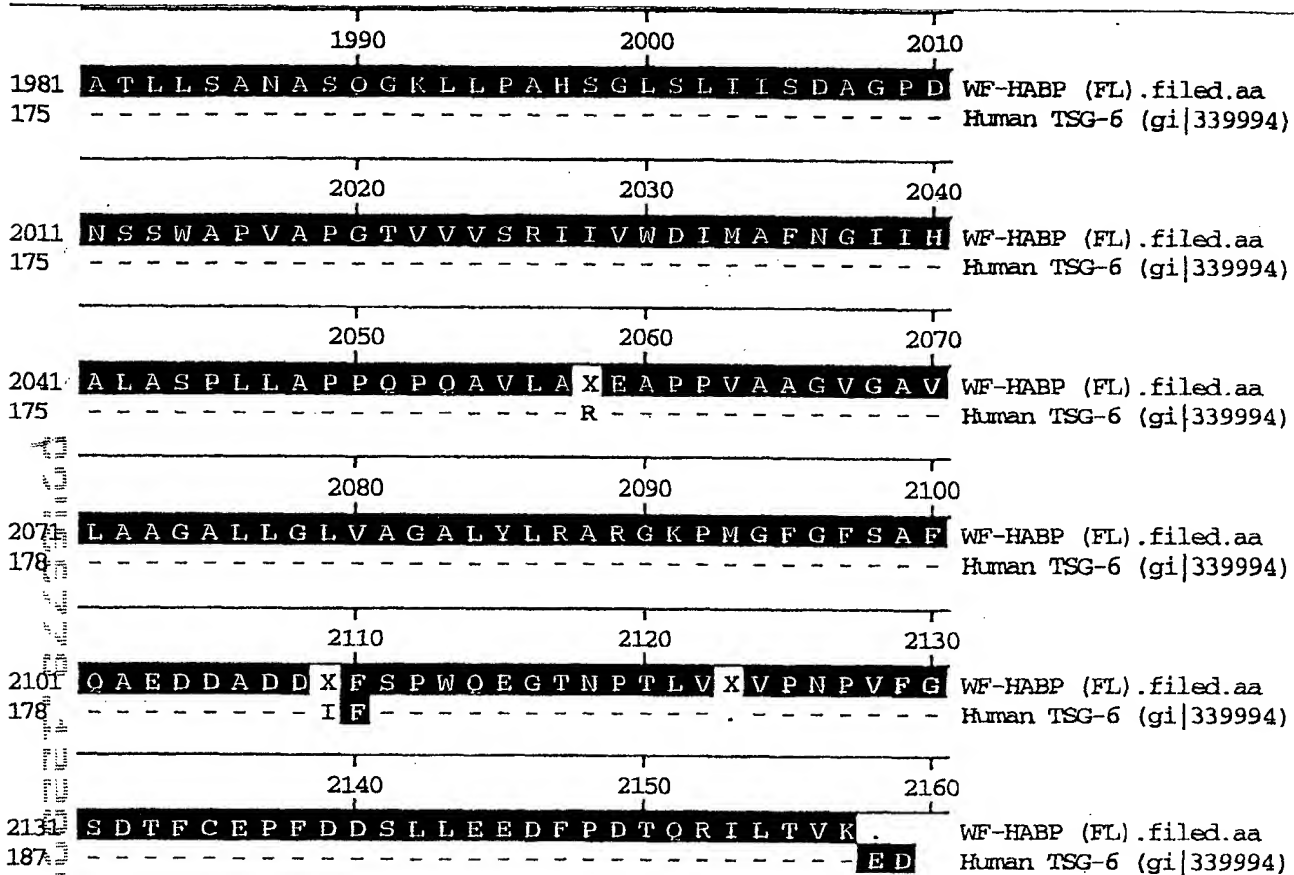


Figure 5G

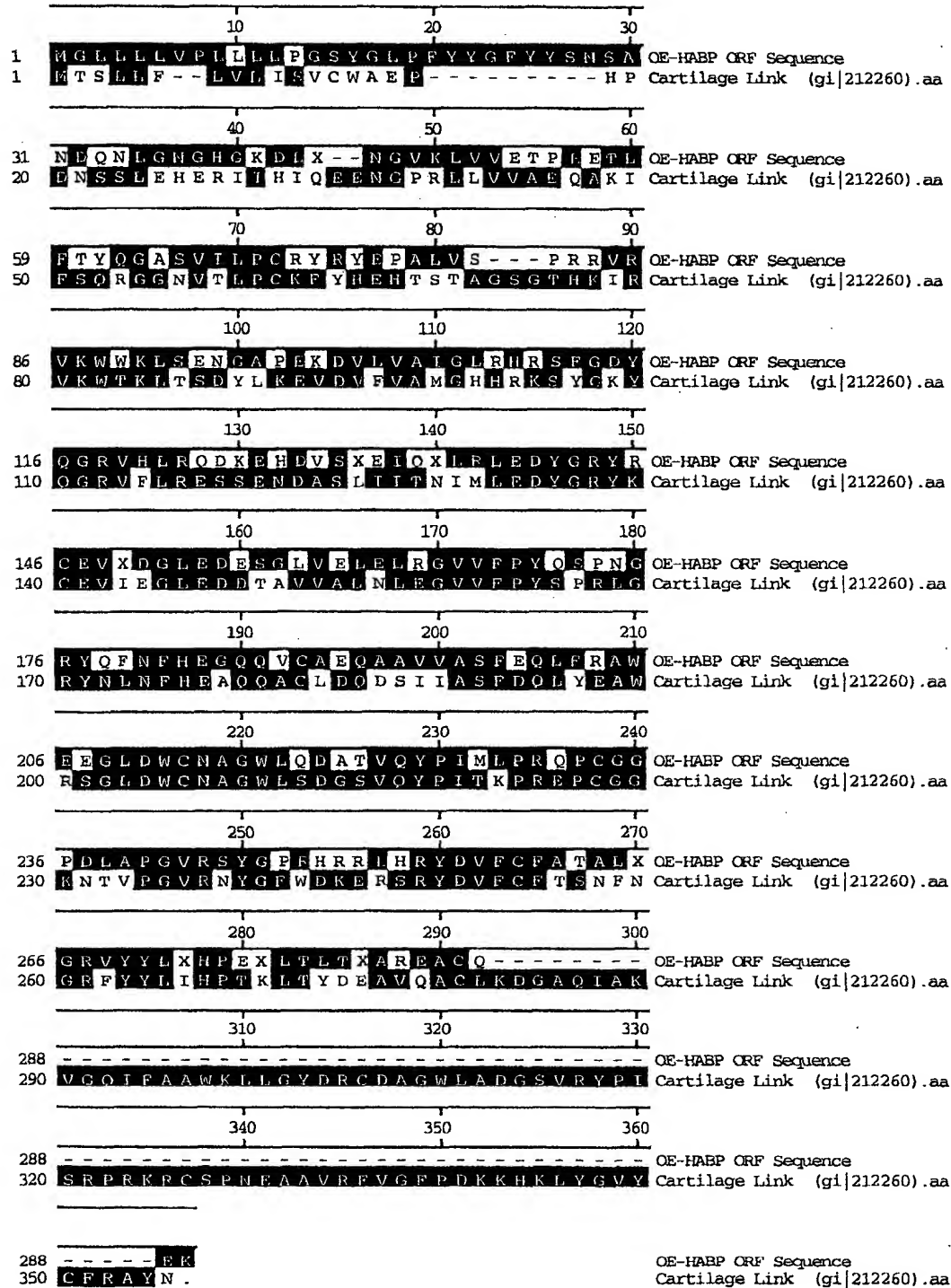


Decoration 'Decoration #1': Shade (with solid black) residues that match the Consensus exactly.



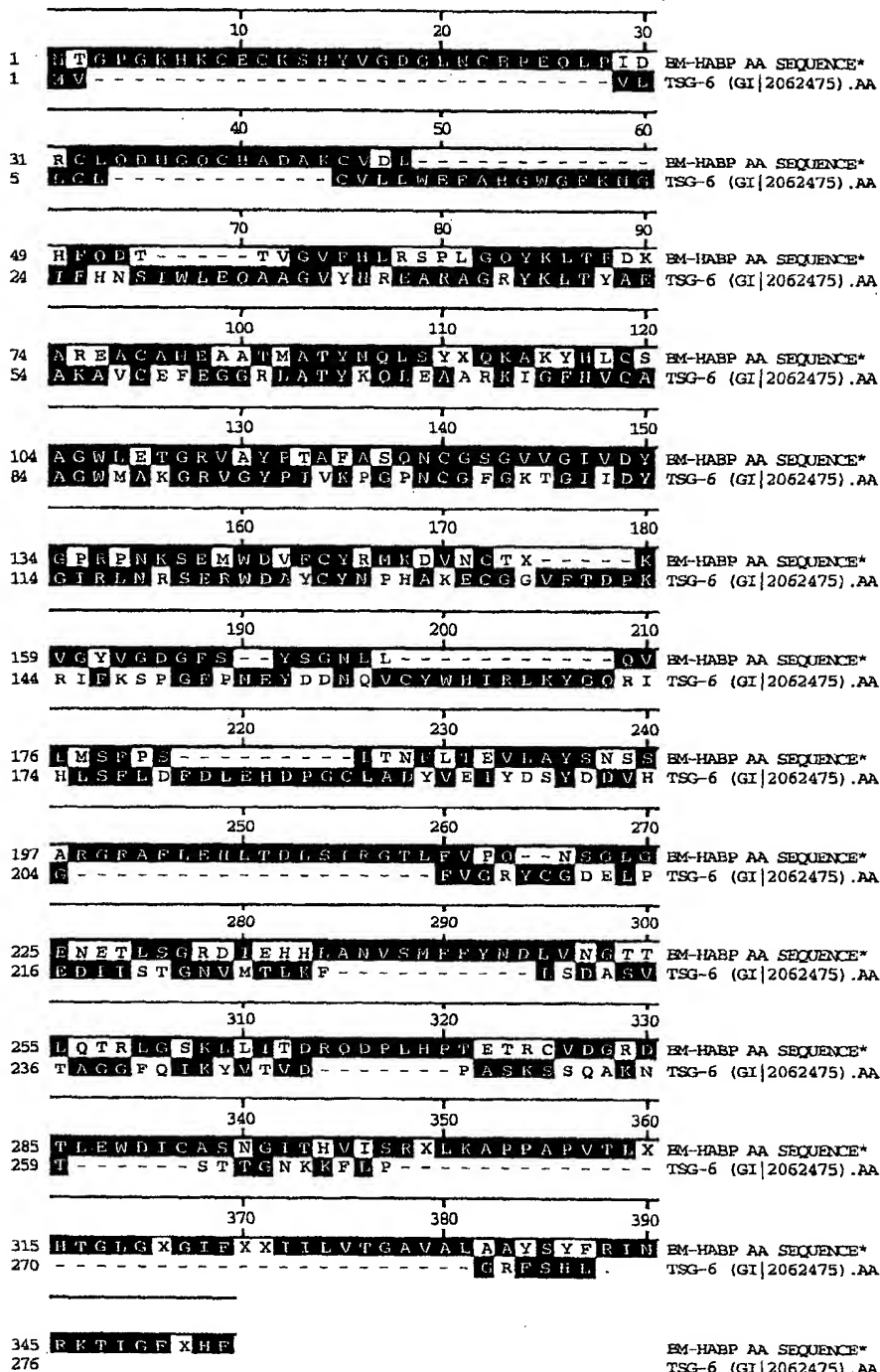
Decoration 'Decoration #1': Shade (with solid black) residues that match the Consensus exactly.

Figure 7



Decoration 'Decoration #1': Shade (with solid black) residues that match the Consensus exactly.

Figure 8



Decoration 'Decoration #1': Shade (with solid black) residues that match the Consensus exactly.

Figure 9

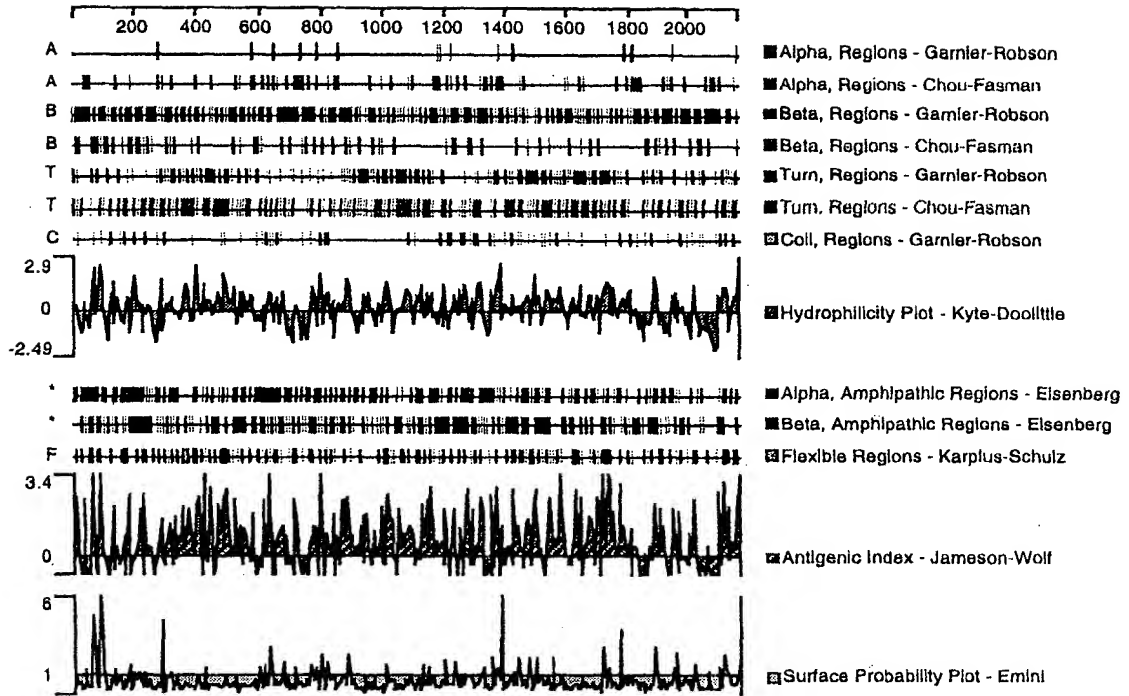


Figure 10

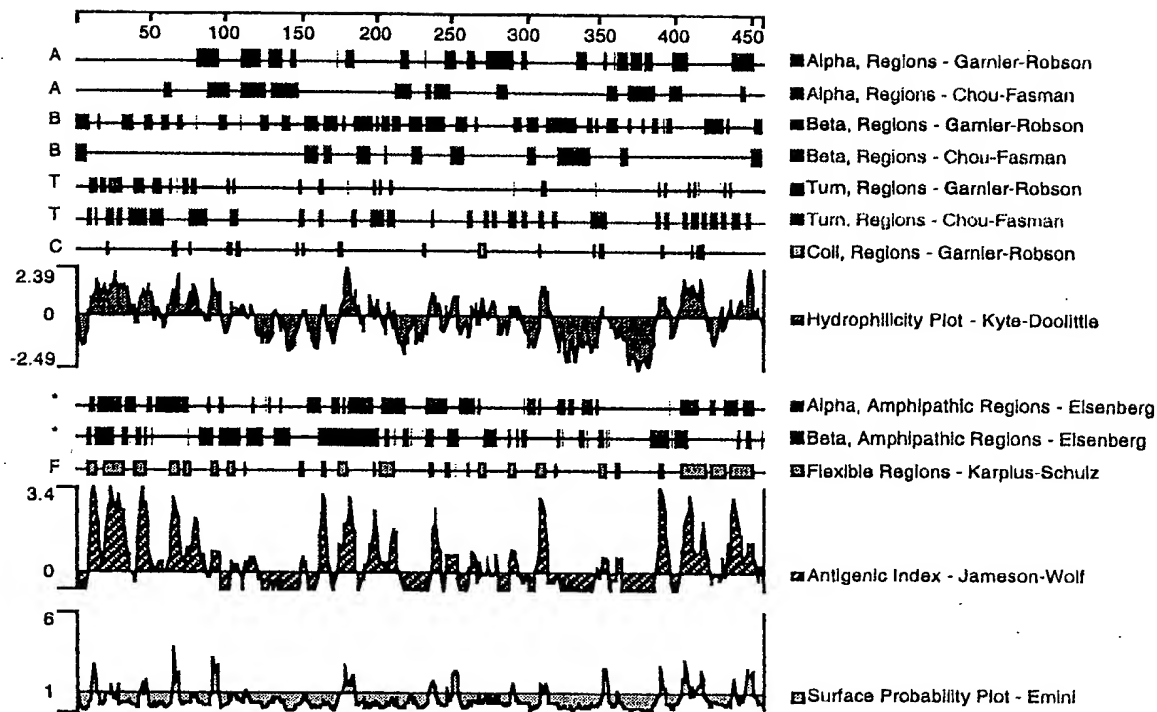


Figure 11

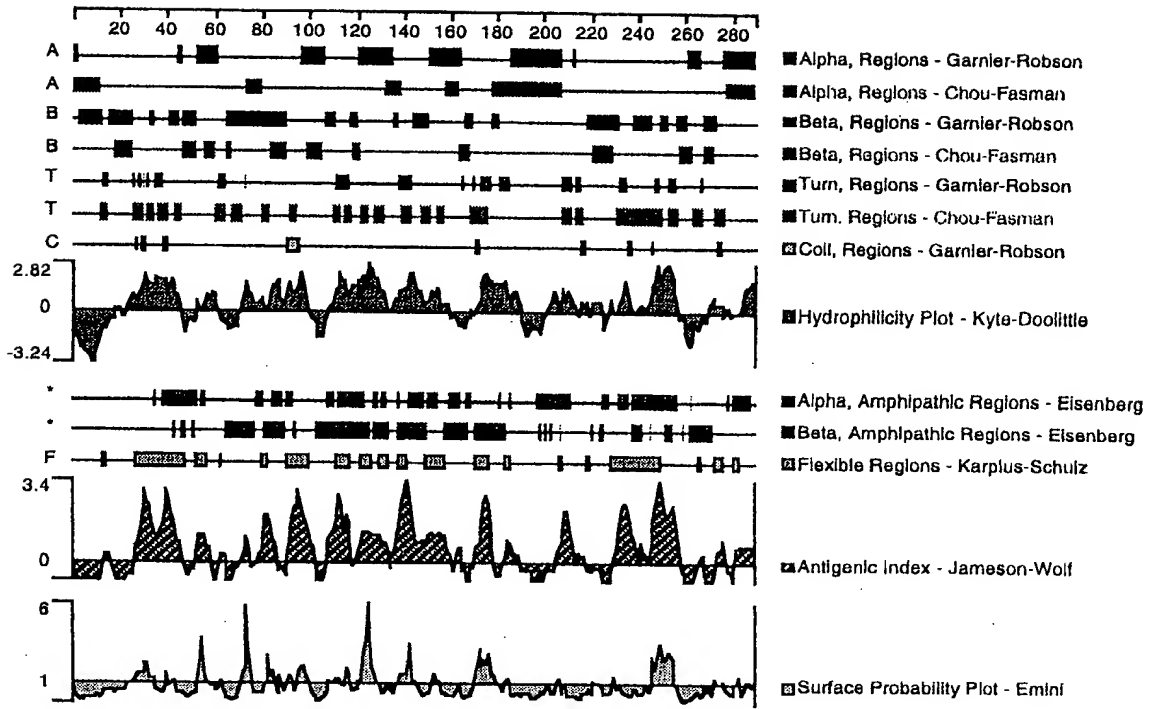


Figure 12

